yang_seonhyeHW30

Question 1

Setting up the data

```r
xvals <- c(1.34, -1.38, -0.19, -0.44, 1.90, -0.80, 0.91, 0.26, 1.37,
           -1.62, -0.96, 1.90, 0.99, 1.96, -1.57, 1.51, -1.61, -1.02, -0.92,
           -1.87, 1.73, -1.23, -1.24, 0.22, 1.42)
yvals <- c(1, 0, 0, 1, 1, 0, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0)
```

Part A

```r
glm_model <- glm(yvals ~ xvals, family=binomial())
summary(glm_model)
```

```
##
## Call:
## glm(formula = yvals ~ xvals, family = binomial())
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.9191  -0.7420   0.4735   0.7469   1.8442
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.1158     0.4865   0.238   0.8118
## xvals         1.0285     0.4060   2.533   0.0113 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 34.617  on 24  degrees of freedom
## Residual deviance: 26.005  on 23  degrees of freedom
## AIC: 30.005
##
## Number of Fisher Scoring iterations: 4
```

We can see from this summary that the parameter estimate for $\hat{\beta}_0 = 0.51441$ and the parameter estimate for $\hat{\beta}_1 = 0.21185$.

Part B

```r
confint(glm_model, 'xvals', level=0.95)
```

```
## Waiting for profiling to be done...
```

```
##      2.5 %    97.5 %
## 0.3154684 1.9571731
```

Looking at these results, we can see that the 95% confidence interval for our $\hat{\beta}_1$ is 0.08406 to 0.33964.

Part C

```r
exp(.1158+1.0285)/(1+exp(.1158+1.0285))
```

```
## [1] 0.7584682
```

We can see there that the probablity of success here is 75.8%

Part D

```r
likelihood.normal = function(theta, x) {
  mu = theta[1]
  sig2 = theta[2]
  n = length(x)
  a1 = (2*pi*sig2)^-(n/2)
  a2 = -1/(2*sig2)
  y = (x-mu)^2
  ans = a1*exp(a2*sum(y))
  return(ans)
}

theta.start = c(.25, .75)

ans <- optim(par=theta.start, fn=likelihood.normal, x=xvals)
mean(ans$par)
```

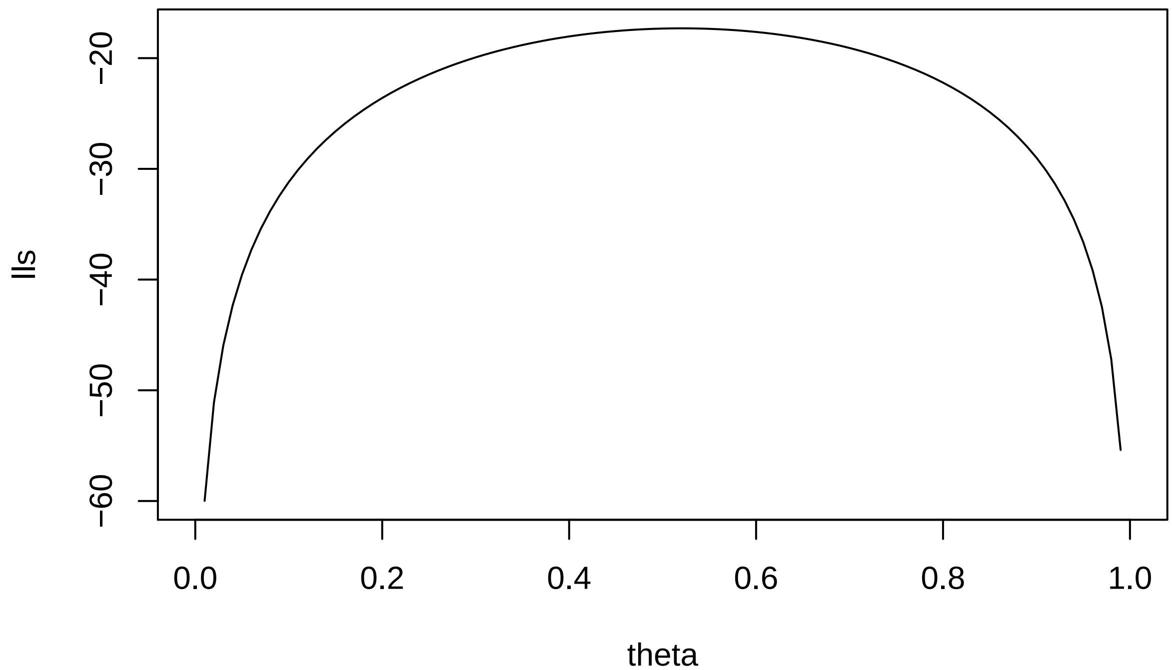
```
## [1] 0.5375
```

We can see here that the value of $\hat{\beta}$ with initial values $(.25, .75)$ is 0.5375.

Part E

```r
log.likelihood <- function(data, theta){
  sum(dbinom(x = data, size = 1, prob = theta, log = T))
}

theta = seq(0, 1, 0.01)
lls <- vector(mode = "numeric", length = length(theta))
for(i in 1:length(theta))
  lls[i] <- log.likelihood(yvals, theta[i])
plot(theta, lls, type = "l")
```




```r
theta[which.max(lls)]
```

```
## [1] 0.52
```

We can see from both the plot and the output values, that our maximum is reached at 0.52, which lines up with the probablity of our input dataset.